

0590
0606

6



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/068,374

DATE: 06/06/2002

TIME: 10:40:49

Input Set : A:\GC652-SEQLIST.txt

Output Set: N:\CRF3\06062002\J068374.raw

ENTERED

```

4 <110> APPLICANT: Shaw, Andrew
6 <120> TITLE OF INVENTION: Modified Target Enzymes Having Catalytic
7   Triads
8   and Uses Thereof
11 <130> FILE REFERENCE: GC652
13 <140> CURRENT APPLICATION NUMBER: US 10/068,374
14 <141> CURRENT FILING DATE: 2002-02-06
16 <150> PRIOR APPLICATION NUMBER: US 60/268,347
17 <151> PRIOR FILING DATE: 2001-02-12
19 <160> NUMBER OF SEQ ID NOS: 5
21 <170> SOFTWARE: FastSEQ for Windows Version 4.0
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 1404
25 <212> TYPE: DNA
26 <213> ORGANISM: Bacillus sp.
28 <400> SEQUENCE: 1
29 atgaaaaaga taactactat ttttgccgta ttgctcatga cattggcggt gttcagtata      60
30 ggaaacacga cagcggtcga tgattattca gttgtagagg aacatgggca actaagtatt      120
31 agtaacggtg aattagtcaa tgaacgaggc gaacaagttc agttaaagg gatgagttcc      180
32 catggtttgc aatggtacgg tcaatttgta aactatgaaa gcatgaaatg gctaagagat      240
33 gattggggaa taactgtatt ccgagcagca atgtatacct cttcaggagg atatattgac      300
34 gatccatcag taaaggaaaa agtaaaagag actgttgagg ctgcgataga ccttggcata      360
35 tatgtgatca ttgattggca taccctttca gacaatgacc cgaatatata taaagaagaa      420
36 gcgaaggatt tctttgatga aatgtcagag ttgtatggag actatccgaa tgtgatatac      480
37 gaaattgcaa atgaaccgaa tggtagtgat gttacgtggg acaatcaaatt aaaaccgtat      540
38 gcagaagaag tgattccggt tattcgtgac aatgacccta ataacattgt tattgtagg      600
39 acaggtacat ggagtcagga tgtccatcat gcagccgata atcagcttgc agatccctaac      660
40 gtcattgtat catttcattt ttatgcagga acacatggac aaaatttacg agaccaagta      720
41 gattatgcac tagatcaagg agcagcgata tttgttagtg aatgggggac aagtgcagct      780
42 acaggtgatg gtggtgtgtt tttagatgaa gcacaagtgt ggattgactt tatggatgaa      840
43 agaaatttaa gctggggcaa ctggtctcta acgcataagg atgagtcac tgcagcgta      900
44 atgccagggt caaatccaac tgggtggttg acagaggctg aactatctcc atctggtaca      960
45 tttgtgaggg aaaaaataag agaatcagca tctattccgc caagcgatcc aacaccgcca      1020
46 tctgatccag gagaaccgga tccaggagaa ccggatccaa cgcacccaag tgatccagga      1080
47 gagtatccag catgggattc aaatcaaatt tacacaaatg aaattgtgta tcataacggt      1140
48 cagttatggc aagcgaaatg gtggacacaa aatcaagagc caggtgaccc atacggtccg      1200
49 tgggaaccac tcaaatctga cccagattca ggagaaccgg atccaacgcc cccaagtgat      1260
50 ccaggagagt atccagcatg ggattcaaat caaatttaca caaatgaaat tgtgtaccat      1320
51 aacggccagc tatggcaagc aaaaatggtg acacaaaatc aagagccagg tgaccatata      1380
52 ggtccgtggg aaccactcaa ttaa                                     1404
54 <210> SEQ ID NO: 2
55 <211> LENGTH: 467
56 <212> TYPE: PRT

```

RAW SEQUENCE LISTING

DATE: 06/06/2002

PATENT APPLICATION: US/10/068,374

TIME: 10:40:49

Input Set : A:\GC652-SEQLIST.txt

Output Set: N:\CRF3\06062002\J068374.raw

```

57 <213> ORGANISM: Bacillus sp.
59 <400> SEQUENCE: 2
60 Met Lys Lys Ile Thr Thr Ile Phe Ala Val Leu Leu Met Thr Leu Ala
61 1 5 10 15
62 Leu Phe Ser Ile Gly Asn Thr Thr Ala Ala Asp Asp Tyr Ser Val Val
63 20 25 30
64 Glu Glu His Gly Gln Leu Ser Ile Ser Asn Gly Glu Leu Val Asn Glu
65 35 40 45
66 Arg Gly Glu Gln Val Gln Leu Lys Gly Met Ser Ser His Gly Leu Gln
67 50 55 60
68 Trp Tyr Gly Gln Phe Val Asn Tyr Glu Ser Met Lys Trp Leu Arg Asp
69 65 70 75 80
70 Asp Trp Gly Ile Thr Val Phe Arg Ala Ala Met Tyr Thr Ser Ser Gly
71 85 90 95
72 Gly Tyr Ile Asp Asp Pro Ser Val Lys Glu Lys Val Lys Glu Thr Val
73 100 105 110
74 Glu Ala Ala Ile Asp Leu Gly Ile Tyr Val Ile Ile Asp Trp His Ile
75 115 120 125
76 Leu Ser Asp Asn Asp Pro Asn Ile Tyr Lys Glu Glu Ala Lys Asp Phe
77 130 135 140
78 Phe Asp Glu Met Ser Glu Leu Tyr Gly Asp Tyr Pro Asn Val Ile Tyr
79 145 150 155 160
80 Glu Ile Ala Asn Glu Pro Asn Gly Ser Asp Val Thr Trp Asp Asn Gln
81 165 170 175
82 Ile Lys Pro Tyr Ala Glu Glu Val Ile Pro Val Ile Arg Asp Asn Asp
83 180 185 190
84 Pro Asn Asn Ile Val Ile Val Gly Thr Gly Thr Trp Ser Gln Asp Val
85 195 200 205
86 His His Ala Ala Asp Asn Gln Leu Ala Asp Pro Asn Val Met Tyr Ala
87 210 215 220
88 Phe His Phe Tyr Ala Gly Thr His Gly Gln Asn Leu Arg Asp Gln Val
89 225 230 235 240
90 Asp Tyr Ala Leu Asp Gln Gly Ala Ala Ile Phe Val Ser Glu Trp Gly
91 245 250 255
92 Thr Ser Ala Ala Thr Gly Asp Gly Gly Val Phe Leu Asp Glu Ala Gln
93 260 265 270
94 Val Trp Ile Asp Phe Met Asp Glu Arg Asn Leu Ser Trp Ala Asn Trp
95 275 280 285
96 Ser Leu Thr His Lys Asp Glu Ser Ser Ala Ala Leu Met Pro Gly Ala
97 290 295 300
98 Asn Pro Thr Gly Gly Trp Thr Glu Ala Glu Leu Ser Pro Ser Gly Thr
99 305 310 315 320
100 Phe Val Arg Glu Lys Ile Arg Glu Ser Ala Ser Ile Pro Pro Ser Asp
101 325 330 335
102 Pro Thr Pro Pro Ser Asp Pro Gly Glu Pro Asp Pro Gly Glu Pro Asp
103 340 345 350
104 Pro Thr Pro Pro Ser Asp Pro Gly Glu Tyr Pro Ala Trp Asp Ser Asn
105 355 360 365
106 Gln Ile Tyr Thr Asn Glu Ile Val Tyr His Asn Gly Gln Leu Trp Gln

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/068,374

DATE: 06/06/2002

TIME: 10:40:49

Input Set : A:\GC652-SEQLIST.txt

Output Set: N:\CRF3\06062002\J068374.raw

```

107      370      375      380
108 Ala Lys Trp Trp Thr Gln Asn Gln Glu Pro Gly Asp Pro Tyr Gly Pro
109 385      390      395      400
110 Trp Glu Pro Leu Lys Ser Asp Pro Asp Ser Gly Glu Pro Asp Pro Thr
111      405      410      415
112 Pro Pro Ser Asp Pro Gly Glu Tyr Pro Ala Trp Asp Ser Asn Gln Ile
113      420      425      430
114 Tyr Thr Asn Glu Ile Val Tyr His Asn Gly Gln Leu Trp Gln Ala Lys
115      435      440      445
116 Trp Trp Thr Gln Asn Gln Glu Pro Gly Asp Pro Tyr Gly Pro Trp Glu
117      450      455      460
118 Pro Leu Asn
119 465
121 <210> SEQ ID NO: 3
122 <211> LENGTH: 35
123 <212> TYPE: DNA
124 <213> ORGANISM: Artificial Sequence
126 <220> FEATURE:
127 <223> OTHER INFORMATION: primer
129 <400> SEQUENCE: 3
130 taaactatat aattgataaa aattttactaa tgaga
132 <210> SEQ ID NO: 4
133 <211> LENGTH: 35
134 <212> TYPE: DNA
135 <213> ORGANISM: Artificial Sequence
137 <220> FEATURE:
138 <223> OTHER INFORMATION: primer
140 <400> SEQUENCE: 4
141 tggcggaata gatgctgatt ctottatattt ttccc
143 <210> SEQ ID NO: 5
144 <211> LENGTH: 308
145 <212> TYPE: PRT
146 <213> ORGANISM: Artificial Sequence
148 <220> FEATURE:
149 <223> OTHER INFORMATION: protein encoded by plasmid pCORE3
151 <400> SEQUENCE: 5
152 Asp Asp Tyr Ser Val Val Glu Glu His Gly Gln Leu Ser Ile Ser Asn
153 1      5      10      15
154 Gly Glu Leu Val Asn Glu Arg Gly Glu Gln Val Gln Leu Lys Gly Met.
155      20      25      30
156 Ser Ser His Gly Leu Gln Trp Tyr Gly Gln Phe Val Asn Tyr Glu Ser
157      35      40      45
158 Met Lys Trp Leu Arg Asp Asp Trp Gly Ile Thr Val Phe Arg Ala Ala
159      50      55      60
160 Met Tyr Thr Ser Ser Gly Gly Tyr Ile Asp Asp Pro Ser Val Lys Glu
161 65      70      75      80
162 Lys Val Lys Glu Thr Val Glu Ala Ala Ile Asp Leu Gly Ile Tyr Val
163      85      90      95
164 Ile Ile Asp Trp His Ile Leu Ser Asp Asn Asp Pro Asn Ile Tyr Lys

```

RAW SEQUENCE LISTING

DATE: 06/06/2002

PATENT APPLICATION: US/10/068,374

TIME: 10:40:49

Input Set : A:\GC652-SEQLIST.txt

Output Set: N:\CRF3\06062002\J068374.raw

```

165          100          105          110
166 Glu Glu Ala Lys Asp Phe Phe Asp Glu Met Ser Glu Leu Tyr Gly Asp
167          115          120          125
168 Tyr Pro Asn Val Ile Tyr Glu Ile Ala Asn Glu Pro Asn Gly Ser Asp
169          130          135          140
170 Val Thr Trp Asp Asn Gln Ile Lys Pro Tyr Ala Glu Glu Val Ile Pro
171 145          150          155          160
172 Val Ile Arg Asp Asn Asp Pro Asn Asn Ile Val Ile Val Gly Thr Gly
173          165          170          175
174 Thr Trp Ser Gln Asp Val His His Ala Ala Asp Asn Gln Leu Ala Asp
175          180          185          190
176 Pro Asn Val Met Tyr Ala Phe His Phe Tyr Ala Gly Thr His Gly Gln
177          195          200          205
178 Asn Leu Arg Asp Gln Val Asp Tyr Ala Leu Asp Gln Gly Ala Ala Ile
179          210          215          220
180 Phe Val Ser Glu Trp Gly Thr Ser Ala Ala Thr Gly Asp Gly Gly Val
181 225          230          235          240
182 Phe Leu Asp Glu Ala Gln Val Trp Ile Asp Phe Met Asp Glu Arg Asn
183          245          250          255
184 Leu Ser Trp Ala Asn Trp Ser Leu Thr His Lys Asp Glu Ser Ser Ala
185          260          265          270
186 Ala Leu Met Pro Gly Ala Asn Pro Thr Gly Gly Trp Thr Glu Ala Glu
187          275          280          285
188 Leu Ser Pro Ser Gly Thr Phe Val Arg Glu Lys Ile Arg Glu Ser Ala
189          290          295          300
190 Ser Ile Pro Pro
191 305

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/068,374

DATE: 06/06/2002

TIME: 10:40:50

Input Set : A:\GC652-SEQLIST.txt

Output Set: N:\CRF3\06062002\J068374.raw